DOCUMENT RESUME

ED 173 798 CS 204 942

AUTHOR Garber, Herbert
TITLE Project Evaluation of English Composition Methods.
PUB DATE Sep 76
NOTE 11p.; Research prepared at State University of New York, College at Oswego

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Freshmen; *Composition (Literary); *Course Evaluation; Higher Education; *Instructional Design; *Large Group Instruction; *Multimedia Instruction; *Small Group Instruction; Teaching Methods; Writing Skills

ABSTRACT

Thirty college students--two sets from each of two freshman writing courses--wrote a total of 120 essays on four assigned topics in a study to measure the effectiveness of large group instruction augmented by multimedia presentations as compared to small group instruction with traditional media presentations. Sixty of the essays were read by four English department professors and rated as excellent, good, fair, or poor. An analysis of the ratings revealed the following: students do improve their expository writing skills in a writing course, the multimedia approach is particularly effective with male freshmen, and the process of judging the merit of undergraduate writing in a dependable, objective manner is extremely difficult and is possibly complicated by assigning topics that may be unfamiliar to some of the writers. An appendix contains a list of the assigned topics. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED173798

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

State University of New York

Research Report

Project Evaluation of English Composition Methods

by
Herbert Garber, Ph.D.
Office of Institutional Research
SUNY, College at Oswego

September, 1976

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Herbert Garber

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS204 942

Project Evaluation of English Composition Methods
(PEECM)

Introduction

A major academic aim at American colleges and universities is to improve undergraduate student writing ability. The College at Oswego has shared this concern for many years.

The English department at this college has explored a number of approaches to facilitating improved expository writing skills. The teaching of composition through the use of multi-media has been one such innovation. Doctor Paul Briand, with support from the Office of Learning Resources, directed by Doctor J. Richard Pfund, has been teaching expository writing for the past six years to large groups of undergraduate students (100 or more per class).

Attempts to improve learning outcomes in any teaching enterprise require objective evaluation if one wishes to know how well the attempt has succeeded. To this end, the present study was conducted in the spring, 1976 semester.

Purpose of Study

As stated in the proposal for this study the principal purpose was:

> To discover how effective and efficient are each of the three methods for teaching English expository writing at the freshman level during the spring 1976 semester.

The proposal, as originally conceived, would have compared the three kinds of basic writing courses then offered: English 102 (Basic Writing for Office of Special Programs Students); English 103 (basic writing, a multi-media approach); English 104 (Writing about Literature). Since English 102 in the spring semester was a second semester of work for the enrolled students it was decided those students were too different from the rest to make comparison meaningful.

Secondary purposes which followed logically from the principal purpose were:

> to discover which student characteristics might interact with which teacher characteristics to facilitate learning expository writing,
>
> to discover how the multi-media technique facilitates learning expository writing as compared to a so-called traditional classroom technique,

to measure the effectiveness of large group instruction augmented by multi-media presentations as compared to a small group instruction with traditional media presentations.

Procedure

1. Subjects

The subjects were two sets of randomly-selected students from the two basic writing courses, English 103 and English 104. Nearly all were freshmen.

To avoid contaminating the results from student knowledge that they were subjects of an experiment, professors of the courses were urged to treat any writing assignments for the study as normal classroom work.

2. Data-gathering: essays

Four English professors selected topics for the subjects to write about. Students wrote in class. The topics were about current social issues and requested the writer to discuss the issue posed (see Appendix A).

Subjects wrote twice during the semester for this evaluation. They wrote one essay during the first week of classes and again, on a different topic, during the last. Approximately thirteen weeks of instruction separated the two writing sessions. Topics were rotated over time and courses so that each topic appeared with about equal frequency for pre-test and post-test sessions and with equal frequency in so-called experimental and control classes. One each of 30 student samples were drawn from the English 103 and 104 courses. One hundred twenty essays were thus selected, xeroxed and later typed for rating.

3. Data-gathering: ratings

Four English department faculty members read and rated sixty essays drawn from the 120 selected. The sixty discards were used for preparing the raters. The first hour of the rating session was spent in preparation for the actual rating process.

After reading a separate, randomly-drawn set of fifteen compositions on a single topic from the discard batch, four essays were chosen as exemplars of excellent, good, fair, and poor freshman expository writing. Ratings of four, three, two, and one, respectively were assigned each essay by each rater independently.

The raters decided, in discussion, that the criteria which most readily distinguished the exemplar essays were: 1) clarity of thesis statement, 2) examples that illustrate

the thesis appropriately, 3) good organization of both the entire essay and individual paragraphs, 4) mechanics, and 5) quality of expression. The criteria were to be weighted in a descending, hierarchical order from clarity of thesis statement regarded as of greatest importance to quality of expression as of least importance. Research, especially at Educational Testing Service, indicates that a rapid, global impressionistic reading leads to greatest inter-rater reliability (McColly and Remstad, 1963, pp. 44, 45).

The sample essays were sorted into four batches of fifteen. Each batch contained essays on a single topic. Also, each batch contained approximately equal proportions of "experimental" and "control", and pre-test and post-test essays. All had been typed exactly as written. Errors of all sorts were retained. All identifying information had been removed. Only a three digit arbitrary identification number appeared on each typed composition. Raters were urged to spend only about one minute reading each essay.

4. Analysis

The essay ratings were analyzed by the Multiple Analysis of Variance (MANOVA) technique. This is a statistical method which enables one to display the individual and joint effects of a number of variables in an experiment.

For the present study the questions that were asked such as about growth in writing ability from taking a course in writing, differences in improvement between two methods of teaching, and differences in such ability change associated with student characteristics such as sex, grade point average, etc. could be answered by the MANOVA approach. Also, the important, although not primary question, of rater reliability is easily answered from this sophisticated analytic technique.

As it happened, rater variability was so great that the planned use of MANOVA yielded results which failed to attain statistical significance. This was so even in the face of clearly discernable differences in growth in writing ability when averages across raters were examined. The fault lay not with the raters as such but rather the circumstances under which they were forced to work.

Therefore, the assumption was made that the <u>average</u> of the four ratings assigned each essay was an acceptable estimate of its true merit. A statistical technique known as the Kruskal-Wallis rank test for independent samples was applied to the gain score for the subjects i.e. the difference between the pre-test average and the post-test average. This seems justified at this time since the rank test statistic is much less powerful than the MANOVA. In other words, if a true difference exists between the abilities estimated in the population (<u>all</u> Oswego freshmen) then the actual

difference in the small sample used in this study would need to be quite large for that true difference to be detected.

Results

A. Student changes

TABLE 1

AVERAGE SCORES ON EACH STUDENT GROUP BY EACH RATER

	PRE-TEST				POST-TEST			
	RATER							
	A	B	C	D	A	B	C	D
Eng. 103 FEML	2.25	2.13	2.25	2.00	2.38	2.38	(2.38)	(2.25)
Eng. 103 MALE	(1.86)	(1.71)	(1.57)	(1.71)	2.00	2.71	2.57	2.43
Eng. 104 FEML	2.00	2.10	2.30	2.00	2.00	2.40	2.40	2.30
Eng. 104 MALE	2.20	2.00	2.40	2.00	(1.60)	(2.00)	2.40	2.40

Table one (above) shows the average rating each rater gave each student group on both post- and pre-tests. The highest average in each rater's column is underlined, the lowest is circled.

It is fairly clear that English 103 attracted male students who were rather poor writers at the beginning of the semester. All four raters agreed about how badly they wrote essays in class. By the end of the semester, these same students were writing better than their classmates who had been significantly better writers 13 weeks earlier. Only one rater (A) felt that the English 103 women were still doing the best writing job at the end of the semester.

Notice also that in every column save one in the right half of the table the English 103 males obtained the highest average ratings that judge assigned to all eight subgroups.

FIGURE 1

CHANGE IN WRITING ABILITY OVER ONE SEMESTER
AVERAGED FOR FOUR RATERS

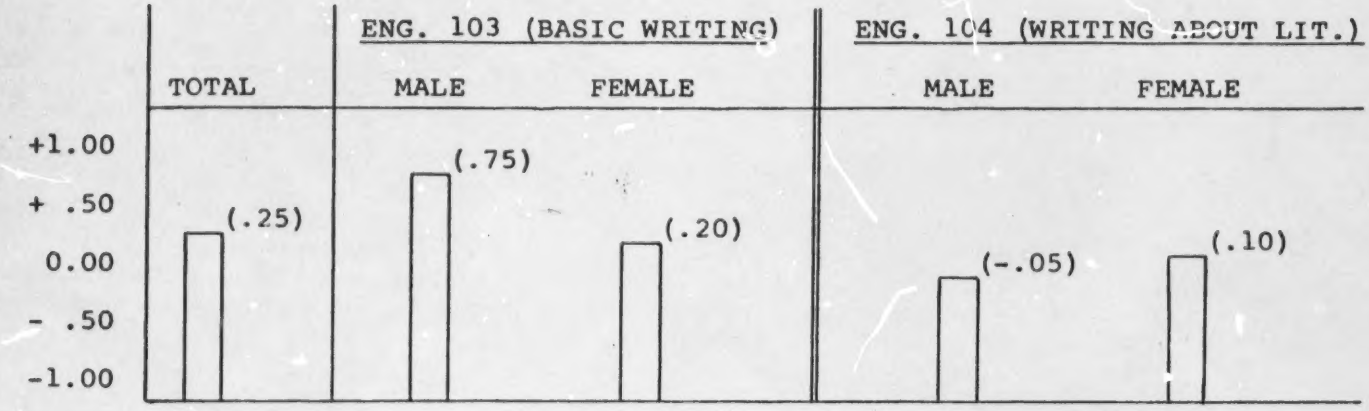

Figure one displays the average change in judged expository writing ability between the beginning and end of the semester. The total change bar shows that overall, the students who took a writing course did indeed improve. The value of .25 means that the average student managed to write a fourth of a rating better by the end of the semester.

The remaining bars show how male and, separately, female students seem to have benefitted in two different courses. This bar graph shows more compactly the changes that were displayed in table one. Male students in the English 103 Basic Writing course advanced, on the average, nearly a full rank (.75) over the semester.

To test the statistical significance of the differences found in these data the rank test for independent samples was calculated. As described earlier, the MANOVA test failed to give the information needed. The rank test result yielded a value of 16.12 which with three degrees of freedom is significant at or beyond the .01 level. That means that the differences among the four groups of students (males, females in two different courses) would have occurred by mere chance less than one time in a hundred.

B. Efficiency

The purpose of the study was twofold. The preceding results seem to show fairly conclusively that students who take a writing course do write better than before they took the course. There is also a differential amount of growth depending on course and sex. The question of effectiveness is thus answered affirmatively.

To answer the question of whether instructional methods

differ in terms of efficiency, we need to examine cost data.

The principal expenditure for college instruction is salary. Since enrollments differ in different courses, raw salary expenditures must be divided by student enrollment figures in order to obtain the usual basis of comparison: unit cost. In the present case, an unusual cost factor must be included. English 103, Basic Writing, required an initial development expenditure for curriculum design, technical assistance, and materials acquisition. Based on the spring, 1976 semester, English 103 cost $134.01 per student while English 104, Writing about Literature, cost $162.65.

Included in the English 103 total unit cost is $50,000 prorated over five years i.e. $10,000 for 1976. That estimate of expenditures for course development was supplied by Doctor Pfund. It includes professorial released time, professional consultation, technicians' salaries, materials, and the current expenditure for technical assistance during classroom presentations. Included also is one-fourth of Doctor Briand's annual salary and the salaries of four graduate assistants.

The English 104 unit cost is comprised only of that portion of the annual salaries of faculty who taught the course attributable to the course itself. It is assumed that any other budgeted expenditures for this more traditional type of instruction are too trivial to include in the analysis.

C. Rater reliability

As mentioned above under Procedure, rater disagreement was quite high.

It is well-known in the field of educational and psychological measurement that to ask human judges to evaluate consistently the abstract properties found in such things as artifacts in art and literature is very difficult. Judges must use internal referents as criteria. It is hard to objectify such criteria.

For this project, the conditions for preparing the panel of judges for their task were less than ideal.

It might be wise for us to rerun the same papers through the rating process with the same or different raters. On this second rating more restricted, better specified instructions must be given the panel.

One other potentially interfering factor deserves mention. Both researcher and raters agreed that the topics assigned the students were troublesome. Current social issues although relatively easy for a mature, fairly

sophisticated writer proved troublesome for a good number of these college freshmen. Future writing tasks for these purposes might better be easy to comprehend, low content demand assignments. Simple, time-worn proverbs seem best.

Conclusion

1. Summary

This study purported to answer two or three questions about the teaching and learning of expository writing among undergraduate students. It was found that:

A. students do improve in such skills from taking a writing course,

B. the multi-media approach is particularly effective with male freshmen,

C. the multi-media approach requires a somewhat lower unit cost for operation, and

D. the process of judging the merit of undergraduate writing in a dependable, objective manner is extremely difficult and is possibly complicated by assigning topics which may be unfamiliar to some of the writers.

2. Further questions

It frequently happens that an evaluative research not only points to answers to the questions it set out to answer but generates new, unexpected problem areas. This study seems to have done that.

Some immediate questions are:

A. To what extent should the College attempt to improve the reliability of essay ratings?

B. Why do female freshmen seem to benefit less in a particular writing course than their male colleagues?

C. Do the gains in writing skills observed in this study transfer to other writing situations (e.g. to other, non-English composition college courses)?

D. How long do the skills demonstrated persist after the course ends?

E. If the acquired skills persist for but a short time what procedures can the College institute which will maintain "good" expository writing performance?

Some long-range questions are:

A. What characteristics of students and courses in writing interact in such a way that improved writing habits will be better facilitated through placing students in the proper writing course(s)?

B. If gains observed in this initial study are found to transfer to only a limited few other settings (see C above) what procedures should the College institute to facilitate a more "general" transfer of good expository writing ability?

C. To what extent does improved writing ability affect positively the after college careers of Oswego students?

APPENDIX A

Topics Assigned to Students

The students in this study were assigned one topic at the beginning of the spring, 1976 semester and a different topic at the end of the same semester. The written instructions and the four topic assignments were:

Topic Ident. A (or B or C or D)

Put your name, topic identification letter, course and section number on cover now. Answer the following question with a legible, organized essay of about 500 words. Please provide a title.

(A) Is attending college necessary to achieve personal fulfillment and professional success? Discuss.

(B) Should smoking be banned in all public areas? Discuss.

(C) The divorce rate has increased drastically in the last fifteen years, is this an indication that marriage is no longer a workable social institution? Discuss.

(D) Some people argue that the excesses of the women's movement have made female chauvinism as prominent as male chauvinism in today's society. Do you agree or disagree? Discuss.

References

McColly, W. and Remstad, R., Comparative Effectiveness of Composition Skills Learning Activities in the Secondary School. Cooperative Research Project 1528, Univ. of Wisconsin, Madison, 1963.